CLAIMS:

1.   A system comprising:

a shared resource;

multiple processors arranged to access said shared resource; and

an operating system configured to allow said multiple processors to perform work on said shared resource concurrently while supporting state changes or updates of said shared resources, said operating system comprising a synchronization algorithm for synchronizing multiple threads of operation with a single thread so as to achieve mutual exclusion between multiple threads performing work on said shared resource and a single thread updating or changing the state of said shared resource without requiring serialization of all threads.

2.   The system as claimed in claim 1, wherein said shared resource includes work queues associated with a hardware adapter configured to send and receive message data to/from a remote system.

3.   The system as claimed in claim 2, wherein said synchronization algorithm is executed to synchronize any thread wishing to update or change a state of said shared resource with all the threads processing I/O operations on said shared resource.

4.   The system as claimed in claim 1, wherein said synchronization algorithm is executed to allow worker threads to work concurrently while processing I/O operations in exclusion of an update thread when a state of said shared resource is not changing, and allow an update thread to change the state or update said shared resource in exclusion of multiple worker threads.

5.   The system as claimed in claim 4, wherein said synchronization algorithm is executed to support a worker thread operation for processing simultaneous I/O operations on said shared resource while concurrently supporting an update thread operation for updating or changing the state of said shared resource.

6.   A system as claimed in claim 5, wherein said worker thread operation is invoked by one of an event and a user's request, and is performed by:

determining whether a lock is available;

if the lock is not available, waiting until the lock becomes available;

if the lock is available, seizing the lock while incrementing a count by a discrete constant to indicate the number of worker threads that are active, and then releasing the lock after the count has been incremented;

after the lock has been released, allowing multiple worker threads to process work concurrently;

1   determining next whether there is work to be processed;

2   if there is work to be processed, processing the work until there is no work to be

3   processed; and

4   if there is no work to be processed, decrementing the count by a discrete constant to

5   indicate when all the worker threads are done with completion processing.

7.   A system as claimed in claim 6, wherein said update thread operation is invoked by

a user's request, and is performed by:

determining whether a lock is available;

if the lock is not available, waiting until the lock becomes available when released by any

one of the worker threads;

if the lock is available, seizing the lock until the count becomes zero (0) to indicate that it

is safe to update or change the state of said shared resource, and updating or changing the state of

8   said shared resource; and

9   after said shared resource has been updated, releasing the lock so as to allow either new

10   worker threads to continue I/O operation processing or a different update thread to continue

11   shared resource updating.

8.   A system as claimed in claim 2, further comprising data channels formed between

2   said system and said remote system, via a switched fabric, and supported by the "*Virtual Interface*

1   *(VI) Architecture Specification"* and the *"Next Generation Input/Output (NGIO) Specification"*

2   for message data transfers between said system and said remote system.

1   9.   The system as claimed in claim 2, wherein said synchronization algorithm is

2   installed as part of a software driver module of an operating system (OS) kernel or an user-level

3   application of said system.

10.   The system as claimed in claim 2, wherein said shared resource includes ones of

work queues, completion queues, FIFO queues, hardware adapters, I/O controllers and other

memory elements of said system.

11.   A network, comprising:

a switched fabric;

3   remote systems attached to said switched fabric; and

4   a host system comprising multiple processors; a host-fabric adapter provided to interface

5   with said switched fabric and included work queues each configured to send and receive message

6   data from a single remote system, via said switched fabric; and an operating system configured to

7   allow said multiple processors to perform work on said work queues concurrently while

8   supporting state changes of said work queues, said operating system comprising a synchronization

9   algorithm for synchronizing multiple threads of operation with a single thread so as to achieve

1   mutual exclusion between multiple threads performing work on said work queues and a single

2   thread changing the state of said work queues without requiring serialization of all threads.

1   12.   The network as claimed in claim 11, wherein said synchronization algorithm is

2   executed to synchronize any thread wishing to update or change a state of said work queues with

3   all the threads processing I/O operations on said work queues.

1   13.   The network as claimed in claim 11, wherein said synchronization algorithm is

2   executed to allow worker threads to work concurrently while processing I/O operations in

3   exclusion of an update thread when the state of said work queues is not changing, and allow an

4   update thread to change the state or update said work queues in exclusion of multiple worker

5   threads.

1   14.   The network as claimed in claim 11, wherein said synchronization algorithm is

2   executed to support a worker thread operation for processing simultaneous I/O operations on said

3   work queues while concurrently supporting an update thread operation for updating or changing

4   the state of said work queues.

1   15.   A network as claimed in claim 14, wherein said worker thread operation is invoked

2   by one of an event and a user's request, and is performed by:

1   determining whether a lock is available;

2   if the lock is not available, waiting until the lock becomes available;

3   if the lock is available, seizing the lock while incrementing a count by a discrete constant

4   to indicate the number of worker threads that are active, and then releasing the lock after the

5   count has been incremented;

6   after the lock has been released, allowing multiple worker threads to process work

concurrently;

determining next whether there is work to be processed;

if there is work to be processed, processing the work until there is no work to be

10   processed; and

11   if there is no work to be processed, decrementing the count by a discrete constant to

12   indicate when all the worker threads are done with completion processing.

1   16.   A network as claimed in claim 14, wherein said update thread operation is invoked

2   by a user's request, and is performed by:

3   determining whether a lock is available;

4   if the lock is not available, waiting until the lock becomes available when released by any

5   one of the worker threads without any work to be processed;

1   if the lock is available, seizing the lock until the count becomes zero (0) to indicate that it

2   is safe to update or change the state of said shared resource, and updating or changing the state of

3   said work queues; and

4   after said work queues have been updated, releasing the lock so as to allow either new

5   worker threads to continue I/O operation processing or a different update thread to continue

6   work queue updating.

17.   A network as claimed in claim 11, further comprising data channels formed

between said host system and said remote systems, via said switched fabric, and supported by the

"*Virtual Interface (VI) Architecture Specification*" and the "*Next Generation Input/Output*

(NGIO) Specification" for message data transfers between said host system and said remote

systems.

18.   The network as claimed in claim 11, wherein said synchronization algorithm is

installed as part of a software driver module of an operating system (OS) kernel or an user-level

application of said host system.

19.   The network as claimed in claim 11, wherein said host system and said remote

systems represent channel endpoints of a data network implemented in compliance with the "*Next*

Generation Input/Output (NGIO) Specification", and data channels formed between said host

28

1   system and said remote systems, via said switched fabric, are supported by the "*Virtual Interface*

2   *(VI) Architecture Specification*" and the "*Next Generation Input/Output (NGIO) Specification*"

3   for message data transfers between said host system and said remote systems.

    20.   A process of synchronizing an update thread which updates a list of work queues

with multiple worker threads which operate on items in the list of work queues in a multi-

processor system, comprising:

    allowing a group of worker threads to concurrently access the list of work queues to

process I/O operations in exclusion of an update thread, when states of the work queues are not

changing;

    incrementing a count of threads processing I/O operations each time a worker thread is

running, while decrementing the count of threads processing I/O operations each time a worker

thread is done processing I/O operations;

10   when the count of threads reaches a designated value indicating that no worker threads are

11   running, allowing an update thread to access and update the list of work queues in exclusion of

12   new worker threads from processing I/O operations; and

13   after the list of work queues is updated, allowing new worker threads to perform I/O

14   operations until all worker threads are done processing I/O operations.

1   21.   A computer-readable medium that stores computer-executable instructions for

2   synchronizing an update thread which updates a list of work queues with a group of threads

3   which operate on items in the list of work queues in a multi-processor system, said computer-

4   executable instructions causing said multi-processor system to:

        permit a group of worker threads to concurrently access the list of work queues to process

I/O operations in exclusion of an update thread, when states of the work queues are not changing;

7   increment a count of threads processing I/O operations each time a worker thread is

8   running, while decrementing the count of threads processing I/O operations each time a worker

9   thread is done processing I/O operations;

10   permit a single update thread to access and update the list of work queues in exclusion of

11   new worker threads from processing I/O operations, when the count of threads reaches a

12   designated value indicating that no worker threads are running; and

13   after the list of work queues is updated, permit new worker threads to perform I/O

14   operations until all worker threads are done processing I/O operations.